

Another proof of undecidability for the correspondence decision problem

Had I been Emil Post

Vesa Halava^{*1,2}

¹Department of Mathematics and Statistics, University of Turku, Finland

²Department of Computer Science, University of Liverpool, UK

November 20, 2014

Abstract

In 1946 Emil Leon Post (*Bulletin of Amer. Math. Soc.* **52** (1946), 264 – 268) defined a famous correspondence decision problem which is nowadays called the Post Correspondence Problem, and he proved that the problem is undecidable. In this article we follow the steps of Post, and give another, simpler and more straightforward proof of the undecidability of the problem using the same source of reduction as Post original did, namely, the Post Normal Systems.

1 Introduction

The original formulation of the *Post Correspondence Problem* (or, as Post called it, *correspondence decision problem*), PCP for short, by Emil Post [4] is the following:

Problem 1 (Post Correspondence Problem). *Let $B = \{a, b\}$ be a binary alphabet, and denote by B^* the set of all finite words over B . Given a finite set of n pairs of words,*

$$W = \{(u_i, v_i) \mid u_i, v_i \in B^*, i = 1, 2, \dots, n\}.$$

Does there exist a nonempty sequence i_1, i_2, \dots, i_k of indices, where each $i_j \in \{1, 2, \dots, n\}$ for $1 \leq j \leq k$, such that

$$u_{i_1} u_{i_2} \cdots u_{i_k} = v_{i_1} v_{i_2} \cdots v_{i_k} ? \tag{1}$$

In the history of computation, the Post Correspondence Problem and its variants have played a major role as a simply defined algorithmically undecidable problems that can be used to prove other undecidability results. Here we concentrate on the undecidability proofs of the PCP itself. In his article [4], Post proved that the problem is unsolvable, or undecidable, as

^{*}vesa.halava@utu.fi

MS Classification: 03D35, 03D03, 68Q01

we say today, by a technical and nontrivial reduction from the *assertion problem of the Post normal systems*. We shall give another proof for the undecidability of the PCP from the same source.

A standard textbook proof of the PCP's undecidability employs the undecidability of the *halting problem of the Turing machines* as the base of reduction, see for example [5], or the construction by Claus [2] from the *word problem of the semi-Thue systems* to the PCP that gives the best known undecidability bounds for n in the definition of the PCP. The number $n = |W|$ of the pairs of words in an *instance* W of the PCP is called the *size* of W .

The standard reduction from the Turing machines or semi-Thue system to the PCP have a common idea: An instance of the PCP is constructed in a way that any solution to it is a (possibly coded) concatenation of all configurations of a required computation or derivation of the original machine or system. This is not the case in Post's original proof of undecidability, indeed, he uses only the words in the rules of an instance of normal system. A sequence of these rule words imply a required derivation in the Normal system, if and only if the sequence is a solution of the instance of the PCP. The new proof presented in this article is based on the idea of a standard type: a solution exists to the constructed instance of the PCP, if and only if the solution is a concatenation of the full configurations required of the given Post normal system.

Finally, in Post's definition the PCP is defined for binary words. Actually, the cardinality of the alphabet B is not relevant, since every instance of the PCP with any alphabet size has an equivalent one in terms of binary words using an injective encoding into binary alphabet $\{a, b\}^*$ from B^* . For example, if $B = \{a_1, a_2, \dots, a_k\}$, then φ defined by $\varphi(a_i) = a^i b$ is such an encoding.

2 Normal systems

We give a formal definition of a normal system instead of the bit informal one used by Post in [4].

Let $A = \{a, b\}$ be a binary alphabet, and let X be a variable ranging over words in A^* . A *normal system* $S = (w, P)$ consists of a *initial word* $w \in A^+$ and a finite set P of *rules* of the form $\alpha X \mapsto X\beta$, where $\alpha, \beta \in A^*$. We say that a word v is a *successor* of a word u , if there is a rule $\alpha X \mapsto X\beta$ in P such that $u = \alpha u'$ and $v = u'\beta$. We denote this by $u \rightarrow v$. Let \rightarrow^* be the reflexive and transitive closure of \rightarrow . Then $u \rightarrow^* v$ holds if and only if $u = v$ or there is a finite sequence of words $u = v_1, v_2, \dots, v_n = v$ such that $v_i \rightarrow v_{i+1}$ for $i = 1, 2, \dots, n - 1$. The Post normal systems are a special case of the *Post canonical systems* for which Post proved in 1943 the Normal-Form Theorem, see [3].

The *assertion* of a normal system $S = (w, P)$ is the set

$$\mathcal{A}_S = \{v \in A^* \mid w \rightarrow^* v\} . \quad (2)$$

The following undecidability result is cited in [4] to Post [3] and for the formal proof there is a reference to Church [1].

Proposition 1. *It is undecidable for a given normal system $S = (w, P)$ and a word $u \in A^+$, whether or not $u \in \mathcal{A}_S$.*

Actually, the problem remains undecidable even if we assume that in each rule $\alpha X \mapsto X\beta$ in P the words α and β are non-empty, and therefore, this is assumed in the following. We shall call the problem, asking for a given word u , whether or not $u \in \mathcal{A}_S$ the *assertion problem*.

3 The proof by Post

The idea of the Post's original undecidability proof is the following: Assume that $u \in \mathcal{A}_S$, where $S = (w, P)$ and let

$$w = \alpha_1 x_1, \quad x_1 \beta_1 = \alpha_2 x_2, \dots, x_{k-1} \beta_{k-1} = \alpha_k x_k, \quad x_k \beta_k = u, \quad (3)$$

where $\alpha_j X \rightarrow X\beta_j$ and $x_j \in A^*$ for all j and $k > 0$. Post proves that existence of a sequence in (3) is equivalent to the following two conditions

$$w\beta_1\beta_2 \cdots \beta_k = \alpha_1\alpha_2 \cdots \alpha_k u \quad (4)$$

and

$$|w\beta_1\beta_2 \cdots \beta_{j-1}| \geq |\alpha_1\alpha_2 \cdots \alpha_j|, \text{ for all } j = 1, \dots, k, \quad (5)$$

where $|v|$ denotes the *length of the word* v . In other words, it is proved that (4) and (5) are equivalent to the condition $u \in \mathcal{A}_S$.

The rest of Post's constructions is the transformation of the system S to a form where the equation (5) holds if (4) holds. Post does this by introducing a new symbol c , considering the reverse words and adding cyclic shifts of all words in \mathcal{A}_S to the assertion of the system. Namely, the normal system $S_1 = (w^R c, P_1)$ where

$$P_1 = \{\alpha^R c X \mapsto X c \beta^R \mid \alpha X \mapsto X \beta \in P\} \cup \{y X \mapsto X y \mid y \in \{a, b, c\}\}$$

is constructed. Next Post proves that $u \in \mathcal{A}_S$ if and only if there are rules $\gamma_j X \mapsto X \delta_j$ in P_1 such that

$$w^R c \delta_1 \delta_2 \cdots \delta_k = \gamma_1 \gamma_2 \cdots \gamma_k u^R c, \quad (6)$$

and that the length condition of the form (5) is true for (6). Indeed, occurrences of the marker symbol c guarantee that the length condition of the

form (5) is satisfied. The reverse words and conjugate rules are added in order to making it possible to work with marked rules.

Finally, Post uses (6) to produce an instance of the PCP. He applies a trick called *desynchronization*; let d be a new symbol and define two mappings ℓ_d and r_d from $\{a, b, c\}^*$ to $\{a, b, c, d\}^*$ such that, for each word $v = a_1 a_2 \cdots a_t$ with $a_i \in \{a, b, c\}$, $\ell_d(w) = da_1 da_2 \cdots da_t$ and $r_d(w) = a_1 da_2 d \cdots a_t d$. Now $u \in \mathcal{A}_S$ if and only if there exists a solution for the instance

$$\{(\ell_d(\delta), r_d(\gamma)) \mid \gamma X \mapsto X\delta \in P_1\} \cup \{(d\ell_d(w^R c), d), (dd, r_d(u^R c)d)\}, \quad (7)$$

of the PCP. Indeed, by desynchronization, a solution to the PCP must begin with $(d\ell_d(w^R c), dd)$, and end with $(dd, r_d(u^R c)d)$. Post concludes, by Proposition 1, that the PCP is undecidable.

4 New proof

As Post, we start with the sequence (3), but use different indices, that is, assume that there exists a sequence

$$w = \alpha_{i_1} x_1, \quad x_1 \beta_{i_1} = \alpha_{i_2} x_2, \dots, x_{k-1} \beta_{i_{k-1}} = \alpha_{i_k} x_k, \quad x_k \beta_{i_k} = u, \quad (8)$$

for a normal system $A = (w, P)$ and input word u where $\alpha_{i_j} X \mapsto X \beta_{i_j} \in P$ for $j = 1, \dots, k$. Instead of equations (4) and (5), we take

$$w x_1 \beta_{i_1} x_2 \beta_{i_2} \cdots x_k \beta_{i_k} = \alpha_{i_1} x_1 \alpha_{i_2} x_2 \cdots \alpha_{i_k} x_k u, \quad (9)$$

where all configurations of the sequence in (8) are concatenated in two ways. Let c and f be new letters and assume that the cardinality of the production set P is t and denote $P = \{p_1, \dots, p_t\}$ where $p_j = \alpha_j X \mapsto X \beta_j$ for $j = 1, \dots, t$. For every $p_j \in P$, we define two pairs of words,

$$p_j^\alpha = (\ell_d(c^j f), r_d(f \alpha_j)) \quad \text{and} \quad p_j^\beta = (\ell_d(\beta_j), r_d(c^j)).$$

where r_d and ℓ_d are the desynchronizing mappings for a new letter d . In other words, we split all productions of P into two pairs. The word $c^j f$ is a marker word forcing us to choose these pairs jointly in a solution of an instance of the PCP defined next. Now, define an instance of the PCP by the pair of words

$$W = \{(d\ell_d(fw), dd), (dd, r_d(fu)d), (da, ad), (db, bd)\} \cup \{p_j^\alpha, p_j^\beta \mid j = 1, \dots, t\}. \quad (10)$$

It is straightforward to prove that $u \in \mathcal{A}_S$ if and only if there exists a solution to the PCP. Indeed, all the solutions to the instance of the PCP are of the

form

$$\begin{aligned}
& dl_d(fwc^{i_1}fx_1\beta_{i_1}c^{i_2}f\cdots c^{i_k}fx_k\beta_{i_k})dd \\
&= dl_d(fw)\ell_d(c^{i_1}f)\ell_d(x_1)\ell_d(\beta_{i_1})\ell_d(c^{i_2}f)\cdots\ell_d(c^{i_k}f)\ell_d(x_k)\ell_d(\beta_{i_k})dd \\
&= ddr_d(f\alpha_{i_1})r_d(x_1)r_d(c^{i_1})r_d(f\alpha_{i_2})r_d(x_2)\cdots r_d(f\alpha_{i_k})r_d(x_k)r_d(c^{i_k})r_d(fu)d \\
&= ddr_d(f\alpha_{i_1}x_1c^{i_1}f\alpha_{i_2}x_2c^{i_2}f\cdots\alpha_{i_k}x_kc^{i_k}fu)d
\end{aligned} \tag{11}$$

implying sequences of the form (8) for the given normal system S .

Finally, note that we are forced to split the rules in two pairs as the words x_i appear in different sides of the words α_i and β_i in (8) and, therefore, α_i and β_i cannot be set in a common pair of words.

5 Conclusion

A new, shorter and bit simpler proof for the undecidability of the PCP was given, using the same source of undecidability, the Post normal systems, as was used in the original proof by Post. Indeed, this new proof could have been found by Post as well, but as a true pioneer of the field of computations he immediately would have noticed the following deficiency of the construction: when considering the size of an instance of the PCP constructed, Post's original construction gives an instance of size $|P| + 5$, but our new construction gives an instance of size $2|P| + 4$. As the undecidable problem in the normal system, the cardinality of P must be at least two, we realize that Post's proof gives a better bound for the undecidability. Therefore, I could not have done anything better - had I been Emil Post.

Acknowledgement. Author thanks Professor Tero Harju for excellent comments.

References

- [1] A. Church, Review of [3], *J. Symb. Logic* **8** (1943), 50 – 52.
- [2] V. Claus, Some remarks on PCP(k) and related problems, *Bull. EATCS* **12** (1980), 54 – 61.
- [3] E. Post, Formal Reductions of the General Combinatorial Decision Problem, *American Journal of Mathematics* **65** (2): 197-215, 1943.
- [4] E. Post, A variant of a recursively unsolvable problem, *Bulletin of Amer. Math. Soc.* **52**, 264 – 268, 1946.
- [5] M. Sipser, *Introduction to the Theory of Computation* (3 ed.), Cengage Learning, 2012.